

退化ポテンシャルを持つ 特異モデルにおける双対平坦構造について

北海道大学大学院 情報科学院 情報理工学コース
中島直道 (Naomichi NAKAJIMA)

概要

甘利, 長岡により導入された双対平坦構造は情報幾何学における中心的な概念であり, 統計科学や機械学習, 最適化問題等における様々な側面に対して見通しの良い統一的な幾何学的解釈をもたらす. しかし, 実応用ではしばしば擬リーマン計量は退化し, 双対平坦構造は定義されない. 本発表では, 接触幾何学, ラグランジュ・ルジャンドル特異点論の観点からそのような特異モデルへの双対平坦構造の一般化を提案し, 甘利, 長岡の拡張ピタゴラスの定理と射影定理がこの特異状況においても成立することを示す.

1 導入

双対平坦構造は情報幾何学における中心的な概念であり, これは, 統計科学, 凸最適化問題, (量子) 情報理論などの様々な領域に統一的な幾何学的視点を与えるものとなっている (甘利・長岡 [1], 甘利 [2], チェンソフ [7]). これはまた, アファイン微分幾何学におけるヘッセ構造と同様の概念である (志磨 [15]). C^∞ -多様体 M において, 双対平坦構造とは三つ組 (h, ∇, ∇^*) のことである. ここで, h は擬リーマン計量 (非退化な対称 $(0, 2)$ -テンソル) であり, ∇ と ∇^* はそれぞれある性質を満たす M 上の平坦アファイン接続である. 双対平坦構造における最も顕著な特徴は, その計量が局所的に ∇ -アファイン座標系におけるあるポテンシャル関数のヘッセ行列で与えられることである. しかしながら, 実応用上では, そのヘッセ行列は M におけるある軌跡 Σ に沿ってしばしば退化するかもしれず, このときには, 厳密に言えば, 微分幾何学的手法が直接的には適用できない. ここでは少し大雑把に, このような空間を特異モデルと呼ぶことにする. 本発表では, 接触幾何学と特異点論の観点から, 特異モデルのあるクラスに対する双対平坦構造の新たな一般化を紹介し, この設定において甘利・長岡の拡張ピタゴラスの定理と射影定理を定式化する (定理 3.10, 3.11). 結果として, 実応用では, これらの定理のほとんどの応用はこのような退化する場合に対しても適切に正当化される.

一般化にあたり次のことに着目した: ポテンシャルのヘッセ行列が退化するとき, 双対ポテンシャル (すなわち, ポテンシャルのルジャンドル変換) のグラフはもはや滑らかな部分多様体にはならないが, これはあるコースティクスに沿って分岐する, 特異点を持った波面 (図 1, 2 を見よ) と呼ばれる対象になる. この双対ポテンシャルのグラフである波面に対し, 接続接束

$$(E', \nabla^{E'}, \Phi' : TM \rightarrow E')$$

を付随させる. ここで, E' は接束 TM の代替物となる M 上のベクトル束であり, その上に平坦接続と h の退化具合を表現する TM からの束写像が備わっている. このとき, 双対平坦空間上の ∇^* の役割は, 特異モデル M 上のベクトル束 E' 上の平坦接続 $\nabla^{E'}$ に取って代わられる. また, この波面の “高さ” を双対ポテンシャルの各値と考えることで, (拡張ピタゴラスの定理などの定式化にも現れる情報幾何学における重要な概念である) 正準ダイバージェンスが特異モデルにおいても一般化される. 元々は, 接続接束の概念は佐治・梅原・山田によって特異波面のリーマン幾何学を研究するために導入され [14], ここではそのアファイン幾何学版を

導入し、構成に用いた。 h が非退化であるとき、従来の双対平坦構造は自然に復元される。

コースティクスと波面の特異点は、ラグランジュ特異点論とルジャンドル特異点論において、古典力学や熱力学、幾何光学、フーリエ積分作用素、制御理論、カタストロフ理論など [3, 4, 5, 6, 9, 10, 13] の幅広い領域の関連の中で徹底的に研究されてきた（アーノルド、ザカリューキン、ヘルマンダーに端を発する）。これらの理論における様々なテクニックや概念を情報幾何学へ持ち込み、両理論の新しい方向性と応用を探る。実際、本研究は純粋数学からデータサイエンスまで異なる背景に由来する様々な関心により動機づけられている。本発表ではその詳細には立ち入らないが、その一つとして、統計的推定や機械学習によるものが挙げられる。実際に、ディープニューラルネットワークを含むほとんどの統計的学習機械ではその計量であるフィッシャー・ラオ行列が退化し [2, 8, 16]、これらに対する新しいアプローチを模索するものとなっている [12]。本稿の残りは次のようになっている。 §2 では、双対平坦構造とその内在的な幾何学を微分幾何学的に定式化した統計多様体の定義とその基本的な性質についてまとめを与える。 §3 では、接触幾何学の基礎事項のまとめと連接接束の定義を与え、退化ポテンシャルを持つ特異モデルに対する双対平坦構造を導入する。最後に、正準ダイバージェンスの一般化を与え、特異モデルにおける拡張ピタゴラスの定理と射影定理の定式化を行う。

本稿を通して、太字は列ベクトルを表すとし (e.g. $\mathbf{x} = (x_1, \dots, x_n)^T$)、プライム記号 \mathbf{x}' は単に \mathbf{x} と区別するために用いる（微分や転置のような操作は意味しない）。また、記号を短くするために、 $\frac{\partial f}{\partial \mathbf{x}}$ で $(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})^T$ を表すとする。簡単のために、多様体と写像のクラスは C^∞ を仮定する。

本発表は、北海道大学の大本享氏との共同研究 [12] に基づく。

2 統計多様体と双対平坦構造

情報幾何学では、統計モデルをリーマン多様体とみなし、そこに双対的なアファイン接続を考えることで様々な領域における問題に統一的な幾何学的視点を与えている。そのような応用を可能にする最も顕著な概念は多様体に備わる双対平坦構造である。統計多様体はこのような双対的構造の内在的な幾何学を微分幾何学の立場から定式化した概念である。本節では、これらの概念の定義と情報幾何学における重要な概念である正準ダイバージェンスについての性質を簡単にまとめる。詳細は [1, 2, 11, 15] を参考にして欲しい。

2.1 統計多様体と双対平坦構造の定義と性質

(M, h) を n 次元擬リーマン多様体、 ∇ を振率がゼロである TM 上のアファイン接続とする。 TM 上のアファイン接続 ∇^* を次で定義する：

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X^* Z).$$

ここで、 X, Y, Z は M 上の任意のベクトル場である。この ∇^* を ∇ の h に関する双対接続と呼ぶ。双対接続は次のような性質を持つ。

命題 2.1. R, R' をそれぞれ ∇, ∇^* に関する曲率テンソルとする。このとき次が成り立つ：

- (i) $R = 0$ なることと $R' = 0$ なることは同値である；
- (ii) $(\nabla^*)^* = \nabla$.

さらに、3 次テンソル $T := \nabla h$ が全対称であるとき、すなわち M 上の任意のベクトル場 X, Y, Z について $(\nabla_X h)(Y, Z) = (\nabla_Y h)(X, Z)$ を満たすとき次が成り立つ：

- (iii) ∇^* の振率はゼロである；
- (iv) $\nabla^{(0)} := \frac{\nabla + \nabla^*}{2}$ は h のレビ・チビタ接続と一致する。

定義 2.2. (統計多様体 [11]) $T = \nabla h$ が全対称であるとき, (M, h, ∇) を統計多様体と呼ぶ.

定義 2.3. (双対平坦空間 [1]) 統計多様体 (M, h, ∇) において ∇ が平坦 ($R = 0$) であるとき, (M, h, ∇, ∇^*) を双対平坦空間と呼び, (h, ∇, ∇^*) を M の双対平坦構造と呼ぶ.

命題 2.1 より, ∇ が平坦であるとき, ∇^* もまたそうであることに注意する.

(M, h, ∇, ∇^*) を双対平坦空間とする. このとき, ∇ に関する局所アフィン座標系 $\mathbf{x} = (x_1, \dots, x_n)$ と ∇^* に関する局所アフィン座標系 $\mathbf{p} = (p_1, \dots, p_n)$ が存在して, 次を満たす:

$$h\left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial p_j}\right) = \delta_{ij} \quad (1 \leq i, j \leq n).$$

このとき, アフィン座標系 \mathbf{p} を h に関する \mathbf{x} の双対座標系と呼ぶ. 簡単のため, アフィン座標系は M 上で大域的に定義されているとする.

命題 2.4. $\mathbf{x} = (x_1, \dots, x_n)$ を ∇ -アフィン座標系, $\mathbf{p} = (p_1, \dots, p_n)$ をその双対座標系とする. このとき, 関数 $f, \varphi \in C^\infty(M)$ が存在して, 次を満たす:

- (i) $\frac{\partial f}{\partial x_i} = p_i, \frac{\partial \varphi}{\partial p_i} = x_i \quad (1 \leq i \leq n),$
- (ii) $f(\mathbf{p}) + \varphi(\mathbf{p}) - \mathbf{x}(\mathbf{p})^T \mathbf{p}(\mathbf{p}) = 0 \quad (\mathbf{p} \in M),$
- (iii) $h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}, h^{ij} = \frac{\partial^2 \varphi}{\partial p_i \partial p_j} \quad (1 \leq i, j \leq n).$

ただし, $[h_{ij}]_{n \times n}, [h^{ij}]_{n \times n}$ はそれぞれ計量 h の座標系 \mathbf{x}, \mathbf{p} に関する成分行列である. このとき f をポテンシャル関数, φ をその双対ポテンシャル関数と呼ぶ.

特に (ii) はルジャンドル双対性と呼ばれ, ポテンシャル関数 f と双対ポテンシャル関数 φ を結びつける:

$$f(\mathbf{p}) + \varphi(\mathbf{p}) - \mathbf{x}(\mathbf{p})^T \mathbf{p}(\mathbf{p}) = 0 \quad (\mathbf{p} \in M).$$

この概念は, 例えば凸解析ではよく現れる概念である; $f(\mathbf{x})$ が凸関数であるとき, $\varphi(\mathbf{p})$ もまたそうであるなどである.

以上の命題 2.4 で示されたように, 双対平坦空間では, ポテンシャル関数のヘッセ行列が計量を与えるという重要な性質がある.

例 2.5. (指数型分布族 [1, 2]) 指数型分布族 M とは, 次の形をした確率密度関数の族である:

$$p_\theta := p(\mathbf{u}|\theta) = \exp(\mathbf{u}^T \theta - \psi(\theta)).$$

ここで, $\mathbf{u} = (u_1, \dots, u_n) \in \mathbb{R}^n$ は確率変数 (その測度は $d\mu$) であり, $\theta = (\theta_1, \dots, \theta_n)^T \in U \subset \mathbb{R}^n$ はパラメータ (U は開集合) である. $M = \{p_\theta\}_{\theta \in U}$ をパラメータ θ により U と同一視し多様体とみなす. $\partial_i := \frac{\partial}{\partial \theta_i}$ とおく. 正規化関数 $\psi(\theta) = \log \int \exp(\mathbf{u}^T \theta) d\mu$ はこの族のポテンシャルと呼ばれる. U のアフィン構造を固定する. このとき, 期待値は対応する双対座標を与える:

$$\eta_i := \mathbf{E}[u_i|\theta] = \int u_i p(\mathbf{u}|\theta) d\mu = \partial_i \psi(\theta).$$

また, (共) 分散

$$h_{ij} := \mathbf{V}[\mathbf{u}|\theta]_{ij} = \partial_i \partial_j \psi(\theta) = \mathbf{E}[(\partial_i \log p_\theta)(\partial_j \log p_\theta)]$$

はフィッシャー情報量によって記述され, これは正定値である. $h = [h_{ij}]$ をリーマン計量, θ, η をそれぞれ ∇, ∇^* -アフィン座標系とみなすとき, (M, h, ∇, ∇^*) は双対平坦空間になる. 正規分布や有限離散分布は典型的な例である.

2.2 正準ダイバージェンス

双対平坦空間 (M, h, ∇, ∇^*) に対して, f, φ をそれぞれ (局所) ポテンシャルとその双対ポテンシャルとする. 簡単のためポテンシャルは大域的に定義されているとする. $M \times M$ 上の関数 $\mathcal{D}: M \times M \rightarrow \mathbb{R}$ を次で定める:

$$\mathcal{D}(p, q) = f(p) + \varphi(q) - \mathbf{x}(p)^T \mathbf{p}(q) \quad (p, q \in M).$$

この \mathcal{D} を M 上の正準ダイバージェンスといい, これは双対構造を考慮した M 上の距離のようなものである. \mathcal{D} の定義はアフィン座標系 \mathbf{x}, \mathbf{p} やポテンシャル f, φ の取り方には依存しないことに注意する. 正準ダイバージェンスは次のような性質を持つ.

命題 2.6. $p \in M$ とする. 正準ダイバージェンス \mathcal{D} に対して, $F := \mathcal{D}(p, -): M \rightarrow \mathbb{R}$ とおくと, 次が成り立つ:

- (i) 任意の $p, q \in M$ に対し $\mathcal{D}(p, q) \geq 0$ であり, $\mathcal{D}(p, q) = 0$ であることと $p = q$ であることは同値である;
- (ii) $\left[\frac{\partial^2 F}{\partial x_i \partial x_j}(p) \right] \left(= \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(p) \right] = [h_{ij}(p)] \right)$ は正定値である.

例 2.7. (指数型分布族の正準ダイバージェンス [1, 2]) 例 2.5 では, $(M = \{p_\theta\}_{\theta \in U}, h, \nabla, \nabla^*)$ は双対平坦空間になり, 正規化関数 $\psi(\theta)$ がそのポテンシャル関数を与えていた. $p_{\theta_1}, p_{\theta_2} \in M$ とする.

$$\begin{aligned} \mathcal{D}(p_{\theta_1}, p_{\theta_2}) &= \psi(\theta_1) + (\theta(p_{\theta_2})^T \eta(p_{\theta_2}) - \psi(\theta_2)) - \theta(p_{\theta_1})^T \eta(p_{\theta_2}), \\ &= \psi(\theta_1) \int p(\mathbf{u}|\theta_2) d\mu + \theta_2^T \int \mathbf{u} p(\mathbf{u}|\theta_2) d\mu - \psi(\theta_2) \int p(\mathbf{u}|\theta_2) d\mu - \theta_1^T \int \mathbf{u} p(\mathbf{u}|\theta_2) d\mu, \\ &= \int (\theta_2^T \mathbf{u} - \psi(\theta_2)) p(\mathbf{u}|\theta_2) d\mu - \int (\theta_1^T \mathbf{u} - \psi(\theta_1)) p(\mathbf{u}|\theta_2) d\mu, \\ &= \int \log(p(\mathbf{u}|\theta_2)) p(\mathbf{u}|\theta_2) d\mu - \int \log(p(\mathbf{u}|\theta_1)) p(\mathbf{u}|\theta_2) d\mu, \\ &= \int p(\mathbf{u}|\theta_2) \log \left(\frac{p(\mathbf{u}|\theta_2)}{p(\mathbf{u}|\theta_1)} \right) d\mu, \\ &:= KL[p_{\theta_2}, p_{\theta_1}]. \end{aligned}$$

この $KL[p_{\theta_2}, p_{\theta_1}]$ はカルバック・ライブラーダイバージェンス (KL ダイバージェンス) と呼ばれ, 情報理論で用いられる確率分布間の差異を測る指標の一つである. つまり, 双対平坦空間 (M, h, ∇, ∇^*) における正準ダイバージェンス \mathcal{D} は KL ダイバージェンスを指数型分布族に制限したものになっている.

3 退化ポテンシャルを持つ特異モデルにおける双対平坦構造

双対平坦空間ではポテンシャル関数のヘッセ行列が計量を与えていた. ヘッセ行列が退化する場合 (リーマン計量が退化する場合) にはどのようなことが起こるだろうか.

(M, ∇) を平坦多様体とし, h を M 上の (退化し得る) $(0, 2)$ -テンソルとする. $\mathbf{x} = (x_1, \dots, x_n)$ を ∇ -アフィン座標系とし, f をその上のポテンシャル関数とする; $h_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. 命題 2.4 によれば f の勾配写像は双対座標系 $\mathbf{p} = (p_1, \dots, p_n)$ への座標変換を与えていた:

$$\frac{\partial f}{\partial x_i} = p_i \quad (1 \leq i \leq n).$$

f のヘッセ行列は勾配写像の微分であるため, ヘッセ行列が退化する場合には, 一般には勾配写像は座標変換を与えない. また, ルジャンドル双対性により双対ポテンシャルのグラフは一般に特異点を持った波面になり,

もはや滑らかな部分多様体にはならない (例 3.1, 3.2 を見よ). 従って, 接束が定義できず双対平坦構造は厳密には破綻する. 本節では, このような退化したポテンシャルを持ち, 退化した計量を許容するような双対平坦構造の自然な一般化を与え, さらに, 情報幾何学における 2 つの主定理 (定理 3.10, 3.11) をこのような特異的設定で定式化する. 以上の定式化は接触幾何学, ラグランジュ・ルジャンドル特異点論に立脚している.

例 3.1. $M = \mathbb{R}^2$ とし, ポテンシャル関数を

$$f(\mathbf{x}) = \frac{x_1^3}{3} + \frac{x_2^2}{2}$$

とする. このとき, $\mathbf{p} = (p_1, p_2) = (x_1^2, x_2)$ であり, f のヘッセ行列 h は $\Sigma = \{x_1 = 0\}$ に沿って退化する.

- ポテンシャル関数のグラフ $W_e(M)$: Σ に沿って変曲点を持つ (図 1 左).
- 双対ポテンシャル関数のグラフ $W_m(M)$: カスピダルエッジを持つ特異曲面である (図 1 右); これは $p_1 \geq 0$ において定義される 2 値双対ポテンシャル

$$z' = \mathbf{p}^T \mathbf{x} - z = \frac{2}{3}x_1^3 + \frac{1}{2}x_2^2 = \pm \frac{2}{3}p_1^{\frac{3}{2}} + \frac{1}{2}p_2^2$$

のグラフであり, $p_1 = 0$ に沿って分岐する.

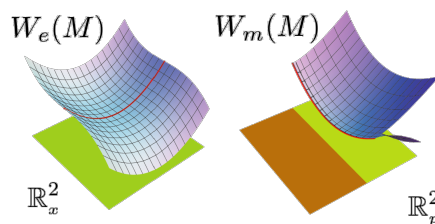


図 1. ポテンシャル関数と双対ポテンシャルのグラフ

例 3.2. $M = \mathbb{R}^2$ とし, ポテンシャル関数を

$$f(\mathbf{x}) = \frac{x_1^4}{4} + \frac{x_2^2}{2}$$

とする. このとき, $\mathbf{p} = (p_1, p_2) = (x_1^3, x_2)$ であり, f のヘッセ行列 h は $\Sigma = \{x_1 = 0\}$ に沿って退化する.

- ポテンシャル関数のグラフ $W_e(M)$: 滑らかであり, 凸である. (図 2 左).
- 双対ポテンシャル関数のグラフ $W_m(M)$: 特異曲面である (図 2 右); これは大域的に定義される双対ポテンシャル

$$z' = \mathbf{p}^T \mathbf{x} - z = \frac{3}{4}x_1^4 + \frac{1}{2}x_2^2 = \frac{3}{4}p_1^{\frac{4}{3}} + \frac{1}{2}p_2^2$$

のグラフであるが, $p_1 = 0$ に沿って特異的である.

3.1 接触多様体

本小節では, 今後用いる接触幾何学の基本的事項の簡単なまとめを行う. N を $2n + 1$ 次元多様体とし, ξ をその上の超平面場とする; $\xi_p \subset T_p N$ ($p \in N$). (N, ξ) が接触多様体であるとは, ξ が $\theta \wedge (d\theta)^n \neq 0$ (非退化

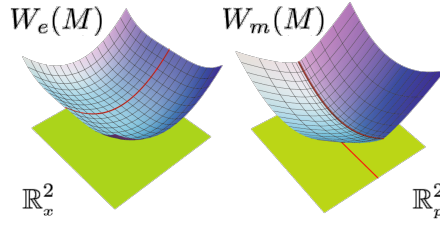


図2. ポテンシャル関数と双対ポテンシャルのグラフ

性) を満たす 1 形式 θ のカーネルによって局所的に表されるときをいう。このとき ξ を N 上の接触構造と呼び、 θ を (局所) 接触形式と呼ぶ。

部分多様体 $L \subset N$ がルジャンドル部分多様体であるとは、 $\dim L = n$ であり、任意の点 $p \in L$ に対し、 $T_p L \subset \xi_p$ なることをいう。接触形式 θ の非退化性により $d\theta$ は超平面場 ξ にシンプレクティック構造を与えるため、一般に、部分多様体 $L \subset N$ が $T_p L \subset \xi_p = \ker \theta_p$ ($p \in L$) を満たすとき、 $\dim L \leq n$ であることに注意する。以下は接触多様体の典型例である。

例 3.3. (標準接触多様体) $N = \mathbb{R}^{2n+1} (= T^*\mathbb{R}^n \times \mathbb{R})$ とし、 $(\mathbf{x}, \mathbf{p}, z)$ を標準的な座標系とする。このとき、

$$\theta = z - \mathbf{p}^T d\mathbf{x} = z - \sum_{i=1}^n p_i dx_i$$

は N 上の接触形式を与え、 $(N, \xi = \ker \theta)$ を標準接触多様体と呼ぶ。この θ は標準接触形式と呼ばれる。

$f : (\mathbb{R}^n, \mathbf{x}) \rightarrow \mathbb{R}$ をその上のポテンシャル関数とする。 (df, f) のグラフ

$$L_f := \left\{ (\mathbf{x}, \mathbf{p}, z) \in \mathbb{R}^{2n+1} \mid p_i = \frac{\partial f}{\partial x_i}, z = f(\mathbf{x}) \right\}$$

は標準接触多様体 (N, ξ) のルジャンドル部分多様体になる；実際、 $df - \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i = 0$ 。

3.2 双対ポテンシャルに付随する連接接束と双対平坦構造

$(\mathbb{R}^{2n+1} = T^*\mathbb{R}^n \times \mathbb{R}, \xi = \ker \theta)$ を標準接触多様体とし、 $(\mathbf{x}, \mathbf{p}, z)$ をその座標系とする。底空間を \mathbb{R}_x^n 、ファイバーを \mathbb{R}_p^n と書き区別することにする。

変換 $\mathcal{L} : (\mathbb{R}^{2n+1}, \xi) \rightarrow (\mathbb{R}^{2n+1}, \xi)$,

$$(\mathbf{x}', \mathbf{p}', z') = \mathcal{L}(\mathbf{x}, \mathbf{p}, z) := (\mathbf{p}, \mathbf{x}, \mathbf{p}^T \mathbf{x} - z)$$

を考えると、これは超平面場を保つ変換である； $\mathcal{L}^* \theta = -\theta$ 。射影 $\pi : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}^n \times \mathbb{R}$, $(\mathbf{x}, \mathbf{p}, z) \mapsto (\mathbf{x}, z)$ との合成を $\pi' := \pi \circ \mathcal{L} : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}_p^n \times \mathbb{R}_{z'}$, $(\mathbf{x}, \mathbf{p}, z) \mapsto (\mathbf{p}, z' = \mathbf{p}^T \mathbf{x} - z)$ とする。

$M = \mathbb{R}_x^n$ とし、 $f : M \rightarrow \mathbb{R}$ とする。 f のヘッセ行列が退化するとき M を退化ポテンシャルを持つ特異モデルと呼ぶことにする。 M と (df, f) のグラフ $L_f \subset (\mathbb{R}^{2n+1}, \xi)$ を同一視する。 π' の L_f への制限を $\pi^m := \pi'|_{L_f} : L_f \rightarrow \mathbb{R}_p^n \times \mathbb{R}_{z'}$, $(\mathbf{x}, \frac{\partial f}{\partial \mathbf{x}}, f(\mathbf{x})) \mapsto (\frac{\partial f}{\partial \mathbf{x}}, \mathbf{x}^T \frac{\partial f}{\partial \mathbf{x}} - f(\mathbf{x}))$ とする。

L_f 上のベクトル束 E' を次で定義する：

$$E' := \{ (p, w) \in L_f \times (\mathbb{R}_p^n \times \mathbb{R}_{z'}) \mid dz'(w) - \mathbf{x}(p)^T d\mathbf{p}(w) = 0 \}.$$

また, L_f がルジャンドル部分多様体であることから, $d\pi^m(T_p L_f) \subset E'_p = d\pi^m(\ker \theta_p)$ ($p \in L_f$) であり, 束写像 $\Phi' := d\pi^m : TL_f \rightarrow E'$ が定まる.

E' 上の接続 $\nabla^{E'}$ を次のように定める. $\tilde{\nabla}'$ を $\mathbb{R}_p^n \times \mathbb{R}_{z'}$ の標準平坦接続とする. また, 各 $p \in L_f$ に対し, $\psi'_p : \mathbb{R}_p^n \times \mathbb{R}_{z'} \rightarrow E'_p$ を z' -軸に沿った線形射影とする.

$$\nabla_X^{E'} s(p) := \psi'_p \circ \tilde{\nabla}'_X s(p)$$

によって接続 $\nabla^{E'}$ を定義する. ここで, s は E' の任意の切断であり, X は L_f の p 周りのベクトル場である.

定義 3.4. (双対ポテンシャルに付随する接続接束 [12]) 組 $(E', \Phi', \nabla^{E'})$ を双対ポテンシャルに付随する接続接束と呼ぶ.

∇ を $M = \mathbb{R}_x^n$ の標準平坦アファイン接続とする.

命題 3.5. $(E', \Phi', \nabla^{E'})$ は次を満たす:

- (i) $\nabla^{E'}$ は平坦である;
- (ii) L_f 上の任意のベクトル場 X, Y に対して, $\nabla_X^{E'}(\Phi'(Y)) - \nabla_Y^{E'}(\Phi'(X)) = \Phi'([X, Y])$ を満たす;
- (iii) ポテンシャル関数 f が非退化なことで, 束写像 $\Phi' : TL_f \rightarrow E'$ が同型であることは同値である;
- (iv) Φ' が同型であるとする. $TL_f (\simeq TM)$ の接続 ∇^* を ∇ の双対接続としたとき, L_f 上の任意のベクトル場 X, Y に対して, $\Phi'(\nabla_X^* Y) = \nabla_X^{E'} \Phi'(Y)$ が成り立つ.

定義 3.6. (退化ポテンシャルを持つ特異モデルに対する双対平坦構造 [12]) $((TL_f, \nabla), (E', \nabla^{E'}))$ を退化ポテンシャルを持つ特異モデル L_f に対する双対平坦構造と呼ぶ.

命題 3.5(iii) より, ポテンシャル関数 f が非退化な場合には, 接続接束 E' は接束 $TL_f \simeq TM$ と同型であり, (iv) より $(E', \nabla^{E'})$ は双対構造 (TL, ∇^*) の自然な一般化であることが分かる. また, このことから, 上で定義した退化ポテンシャルを持つ特異モデルに対する双対平坦構造もまた, 定義 2.3 の意味で, 従来の双対平坦構造の自然な一般化になっている.

3.3 拡張ピタゴラスの定理と射影定理

$f : M \rightarrow \mathbb{R}$ を $M = \mathbb{R}_x^n$ のポテンシャル関数とし, $L_f \subset (\mathbb{R}^{2n+1}, \xi = \ker \theta)$ を (df, f) のグラフとする. $((TL_f, \nabla), (E', \nabla^{E'}))$ を L_f の双対平坦構造とする. $c : I \rightarrow L_f$ を曲線とする. ここで $I (\neq \emptyset) \subset \mathbb{R}$ は開区間であり, 速度ベクトル場を $\dot{c}(t) := \frac{d}{dt} c(t) \in T_{c(t)} L_f$ ($t \in I$) とする.

定義 3.7. (m -曲線 [12]) 曲線 $c : I \rightarrow L_f$ が m -曲線であるとは, はめ込み $(\dot{c}(t) \neq 0)$ であり, 任意の $t \in I$ に対し, $E'_{c(t)}$ のベクトル

$$\Phi' \circ \dot{c}(t), \nabla_{\dot{c}}^{E'}(\Phi' \circ \dot{c})(t), (\nabla_{\dot{c}}^{E'})^2(\Phi' \circ \dot{c})(t), \dots$$

が同時にはゼロにならず, 任意の 2 つが線形従属であるときをいう.

ここで, 情報幾何学の慣習に従い, \mathbb{R}_p^n 側における概念に記法 “ m -” を用いた. また, \mathbb{R}_x^n 側における概念には “ e -” をつけて表すことにする.

m -曲線 $c(t)$ の条件から, $\pi^m(c(t))$ は \mathbb{R}_p^n における同一直線上を動き, これは m -測地線 ($\nabla^* = \nabla^{E'}$ に関する測地線) の再パラメータ化である. ただし, ポテンシャルが退化するとき, 勾配写像の特異値集合に触れると, $\pi^m(c(t))$ は特異値集合上で速度をゼロとして, 同一直線上を引き返す (図 3 左) かあるいはそのまま直進

する (図 3 右). いずれの場合においても, $\pi^m(c(t))$ はある直線上を動くことに注意する. この直線の方向ベクトルを \mathbf{m} とする. e -測地線 (∇ に関する測地線) に関してはポテンシャルの退化性に関わらず, 従来と同様に扱えることに注意する. e -測地線が定める \mathbb{R}_p^n における直線の方向ベクトルを \mathbf{e} とする.

以下の定義は方向ベクトルの取り方によらない.

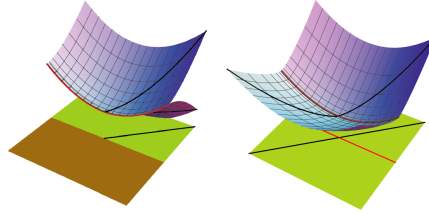


図 3. \mathbb{R}_p^n への射影 π^m による m -曲線の像

定義 3.8. c_e, c_m をそれぞれ, e -測地線, m -曲線とする. \mathbf{e}, \mathbf{m} をそれぞれ上記の意味の方向ベクトルとする. S を L_f の部分多様体とし, $q \in S$ において c_m は S に交わるとする. c_m と S が q において直交するとは, 任意の $u \in T_q S$ に対し, 次が成り立つときをいう:

$$\mathbf{m}^T d\mathbf{x}(u) = 0.$$

同様に, c_e と S の q における直交性は, 任意の u に対して $\mathbf{e}^T d\mathbf{p}(u) = 0$ が成り立つときをいう. さらに, c_e と c_m が q において真に直交しているとは, $\mathbf{e}^T \mathbf{m} = 0$ が成り立つときをいう.

定義 3.9. 正準ダイバージェンス $\mathcal{D} = \mathcal{D}_{L_f} : L_f \times L_f \rightarrow \mathbb{R}$ を次で定義する:

$$\mathcal{D}(p, q) = z(p) + z'(q) - \mathbf{x}(p)^T \mathbf{p}(q) \quad (p, q \in L_f).$$

ここで, $z(p) = f(\mathbf{x}(p)), z'(q) = \mathbf{x}(q)^T \mathbf{p}(q) - z(q)$ である.

定理 3.10. (拡張ピタゴラスの定理 [12]) $p, q, r \in L_f$ を 3 つの異なる点とし, p と q は e -測地線 c_e によって繋がれ, q と r は m -曲線 c_m によって繋がれているとする. さらに, c_e と c_m が q において真に直交しているとする. このとき, 次が成り立つ:

$$\mathcal{D}(p, r) = \mathcal{D}(p, q) + \mathcal{D}(q, r).$$

定理 3.11. (射影定理 [12]) S を L_f の部分多様体とし, $c_m : [0, 1] \rightarrow L_f$ を $q = c_m(1) \in S$ なる m -曲線とする. $p = c_m(0) \in L_f$ とおく. このとき, c_m と S が q において直交することと, q が関数 $F = \mathcal{D}(-, p) : S \rightarrow \mathbb{R}$ の臨界点になることは同値である. e -測地線 c_e と $F = \mathcal{D}(p, -)$ に対しても同様に成り立つ.

参考文献

- [1] S. Amari, H. Nagaoka, *Method of information geometry*, A.M.S., Oxford Univ. Press (2000).
- [2] S. Amari, *Information Geometry and Its Application*, Applied Math. Sci., 194, Springer (2016).
- [3] V.I. Arnol'd, *Mathematical Methods of Classical Mechanics*, 2nd Edition, Grad. Texts Math. **60**, Springer-Verlag (1989).
- [4] V.I. Arnol'd, S.M. Gusein-Zade, A.N. Varchenko, *Singularities of Differentiable Maps I*, Monographs in Math. **82**, Birkhäuser (1986).

- [5] V.I. Arnol'd, *Singularities of Caustics and Wave Fronts*, Kluwer Acad. Publ. (1990).
- [6] V.I. Arnol'd, *Catastrophe Theory*, 3rd Edition, Springer-Verlag (1992).
- [7] N. Chentsov, *Statistical decision rules and optimal inference*, Translation of Math. Monograph **53**, AMS, Providence (1982).
- [8] K. Fukumizu and S. Kuriki, *Statistics of Singular Models*, Frontier Stat. Sci. 7, Iwanami Publ. (2004) (in Japanese).
- [9] S. Izumiya and G. Ishikawa, *Applied Singularity Theory* (in Japanese), Kyoritsu Shuppan, Co. Ltd. (1998)
- [10] S. Izumiya, M.C. Romero-Fuster, M.A.S. Ruas, F. Tari, *Differential Geometry from a Singularity Theory Viewpoint*, World Scientific (2015).
- [11] H. Matsuzoe, Statistical manifolds and affine differential geometry, *Advanced Stud. Pure Math.* **57** (2010), 303–321.
- [12] N. Nakajima, T. Ohmoto, *The dually flat structure for singular models*, preprint, arXiv:2011.11219 (2020).
- [13] T. Poston, I. Stewart, *Catastrophe Theory and Its Applications*, Pitman Publ. Ltd. (1978).
- [14] K. Saji, M. Umehara, K. Yamada, The geometry of fronts, *Annals of Math.*, **169** (2009), 491–529.
- [15] H. Shima, *The geometry of Hessian Structures*, World Scientific (2007).
- [16] S. Watanabe, *Algebraic Geometry and Statistical Learning Theory*, Cambridge Univ. Press (2008).